DOCUMENT RESUME

ED 346 097 SP 033 869

AUTHOR	Raudenbush, Stephen W.; And Others
TITLE	Teaching for Higher-Order Thinking in Secondary Schools: Effects of Curriculum, Teacher Preparation, and School Organization.
INSTITUTION	Center for Research on the Context of Secondary School Teaching.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	CRC-P92-146
PUB DATE	Jun 92
CONTRACT	OERI-G0087C235
NOTE	53p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Educational Objectives; English; Policy Formation; Research Needs; *School Organization; Secondary Education; *Secondary School Curriculum; Secondary School Mathematics; Secondary School Science; Secondary School Students; Secondary School Teachers; Social Studies; Teacher Education; *Teaching Experience; *Teaching Skills; *Thinking Skills; Track System (Education)
IDENTIFIERS	*Higher Order Learning; *Knowledge Base for Teaching

ABSTRACT

This study was conducted to explore variations in emphasis on teaching for higher-order thinking in secondary schools. Three explanations are considered: (1) conceptions of teaching and learning encourage teachers to pursue higher-order objectives primarily when teaching high-track students and advanced courses; (2) lack of adequate preparation to teach higher-order thinking; and (3) aspects of school organization discouraging teachers from pursuing higher-order objectives. Secondary teachers in 16 schools were asked to identify instructional goals for each of their classes. Scales, intended to capture the emphasis on higher-order objectives in math, science, social studies, and English, were constructed and evaluated. Findings revealed that math classes differ in instructional objectives because math teachers tend to differentiate objectives for different classes, while social studies classes vary primarily because of differences among teachers; school-to-school variation was exhibited in English and social studies, but not in math and science; the effect of tracking on higher-order objectives is powerful in all disciplines, but especially in math and science; and effects of teacher preparation and organization design were manifest in English and social studies, but not in math or science. Implications for policy and future research are considered. (Author/LL)

ED346097

CRC

Center for Research on the Context of Secondary Teaching
School of Education, CERAS Building, Stanford University, Stanford, CA 94305-3084

TEACHING FOR HIGHER-ORDER THINKING IN SECONDARY SCHOOLS: EFFECTS OF CURRICULUM, TEACHER PREPARATION, AND SCHOOL ORGANIZATION

P92-146

Stephen W. Raudenbush, Michigan State University
Brian Rowan, University of Michigan
Yuk Fai Cheong, Michigan State University

TEACHING FOR HIGHER-ORDER THINKING IN SECONDARY SCHOOLS: EFFECTS OF CURRICULUM, TEACHER PREPARATION, AND SCHOOL ORGANIZATION

P92-146

Stephen W. Raudenbush, Michigan State University
Brian Rowan, University of Michigan
Yuk Fai Cheong, Michigan State University

June 1992

Research reported here was conducted under the auspices of the Center for Research on the Context of Secondary School Teaching with funding from the Office of Educational Research and Improvement, U.S. Department of Education, Cooperative Agreement #OERI-G0087C235. We wish to thank our colleagues Milbrey McLaughlin and Joan Talbert for their crucial contribution in developing the questionnaire utilized in our study and Marian Eaton for her careful preparation of the data.

Abstract

We consider three explanations for variation in emphasis on teaching for higher-order thinking in US secondary classrooms: 1) conceptions of teaching and learning rooted in the secondary school curriculum encourage teachers to pursue higher-order objectives primarily when teaching high-track students and advanced courses; 2) many teachers lack adequate preparation to teach for higher-order thinking; and 3) aspects of school organization discourage teachers from pursuing higher-order objectives. To test hypotheses flowing from these explanations, we asked secondary teachers in 16 schools to identify their instructional goals for each of their classes. From their responses, we constructed and evaluated scales intended to capture the emphasis on higher-order objectives in math, science, social studies, and English. A variance decomposition revealed that math classes differ in instructional objectives primarily because each math teacher tends to differentiate her objectives for her different classes, while social studies classes vary primarily because of differences among social studies teachers. School-to-school variation was manifest in English and social studies but not in math and science. A three-level hierarchical regression model revealed powerful effects of track on higher-order objectives in all disciplines, but especially in math and science. Effects of teacher preparation and organization design were manifest in English and social studies, but not in math or science. The differentiation of instructional objectives based on student track is apparently deeply institutionalized in all subjects, but particularly in math and science. Implications for policy and future research are considered.

Over the past decade, research on teaching has undergone an important transformation as researchers have turned their attention from instruction for basic skills toward teaching for higher-order thinking (Bereiter and Scardamalia, 1987; Peterson, 1988; Prawat, 1989; Newmann, 1990). This change results from a widespread perception that US schools are comparatively ineffective in cultivating conceptual understanding of academic subjects. For example, recent large-scale assessments demonstrate that although US students perform adequately on tests of basic skills, they perform comparatively poorly on tasks that involve problem solving, critical analysis, and flexible understanding of subject matter (US Department of Education, 1991: 32-41). In this light, teaching for higher-order thinking is increasingly accepted as a goal toward which the education profession should strive.

Recent commentary and research also emphasize that teaching for higher-order thinking is important for the learning of <u>all</u> students in <u>all</u> academic subjects. For example, Peterson (1988:2) cites research suggesting "the need for an increased instructional focus on teaching higher-level skills in mathematics to all students. Such an increased focus might be particularly important for lower-achieving students, who have more difficulty than their peers in learning these higher-order skills on their own." With respect to secondary social studies, Newmann (1990:48) proposed that teaching for higher-order thinking is important for all learners: "...Any person, young or old, regardless of experience, can participate in higher-order thought. Students will differ in the kinds of challenges they are able to master, but all are capable of confronting a challenge in the interpretation, analysis, and manipulation of knowledge." Doyle (1983) also reviews research indicating that teaching cognitive processes and knowledge structures can be of special benefit to lower-achieving students.

Despite the emerging consensus on the importance of teaching for higher-order thinking, research generally finds that classroom instruction in high schools is focused on basic skills (Goodlad, 1984; Powell, Cohen, & Farrar, 1985). To the extent that teaching for higher-order thinking is manifest, evidence suggests that it occurs far more often in high-track than in low-track classes (Metz, 1978; Oakes, 1985; Page, 1990). Thus, at the high school level in the United States, a sharp contrast exists between current visions of educational excellence and currently institutionalized patterns of educational practice.

In this paper, we consider and test several explanations for the apparent disparities across US secondary classrooms in teaching for higher-order thinking. The first explanation holds that hierarchical conceptions of teaching and learning are deeply embedded in the high school curriculum and that these conceptions discourage teachers from embracing higher-order instructional goals except when teaching the most advanced students in the most advanced subjects. If this explanation is credible, we should expect instructional goals to vary <u>within</u> teachers. That is, when the several classes taught by a given high school teacher vary in terms of the presumed academic ability of their students or the level of subject matter to be taught, that teacher should place more emphasis on teaching for higher-order thinking when teaching high-track students and when teaching more advanced subjects. Moreover, this tendency of a teacher to differentiate instructional goals across classes should be most pronounced in those academic disciplines that are most clearly differentiated on the basis of student ability and level of course difficulty.

A second explanation emphasizes differences among teachers rather than differences within teachers. According to this view, variations in teachers' subject-matter knowledge and pedagogical expertise are critical in understanding differences among teachers in instructional goals (Shulman, 1987). We hypothesize that more highly educated and experienced teachers, who are presumably more knowledgeable about their discipline and have acquired more pedagogical expertise, will embrace higher-order instructional goals with higher frequency than will less educated or less experienced teachers. We also investigate the effect of a teacher's preparation to teach the specific content of each class to which the teacher is assigned. We hypothesize that when teacher preparation matches the content, the teacher will be more likely to pursue higher-order instructional goals than when preparation and content match poorly. The adequacy of the match will depend both on the teacher's overall preparation and on the process used to assign teachers to courses.

A third explanation focuses on the organizational environments in which teachers work. Recent work on school restructuring, for example, suggests that bureaucratic rules that closely regulate or routinize teachers' work will discourage the pursuit of higher-order instructional goals while organizational environments characterized by supportive administrative leadership, high levels of teacher collaboration, and strong teacher control over instruction will facilitate the pursuit of such instructional goals. According to reasoning described in detail later, organizational environments can be expected to produce variation in instructional goals at two levels: among the teachers within a school and among schools.

Using teachers' self-reported emphasis on teaching for higher-order goals in each of the classes they teach, we examine the incidence of teaching for higher-order thinking in a sample of 1205 classes taught by 303 academic teachers in 16 diverse high schools in California and Michigan. The three explanations we consider imply that teachers' goals will vary at each of three levels: within teachers as a function of characteristics of the classes a teacher encounters; among the teachers working in a particular school; and across schools. Accordingly, we employ a three-level hierarchical linear model that enables us to decompose variation in instructional goals into these three components. We then employ predictor variables measured at each of these three levels in an attempt to account for the manifest variation at each level. This procedure is replicated for each of four major academic disciplines: mathematics, science, social studies, and English.

We begin by reviewing the logic underlying the three explanations for variation in higher-order instructional goals. We then describe in detail the methods we use to test these explanations, including the scales used to measure emphasis on higher-order instructional goals in each discipline. Finally, we turn to the results and implications for research on teaching and school reform.

Alternative Explanations and Hypotheses

Conceptions of Learning Rooted in the Secondary School Curriculum

A number of studies have reported that classes serving high-achieving students are substantially more likely than classes serving low-achieving students to emphasize higher-order thinking processes (Hargreaves, 1967; Metz, 1978; Oakes, 1985; Page 1990). A picture emerging from qualitative accounts is that educators commonly presume less able students to he capable primarily of rote memorization and procedural knowledge. In this view, more able students can apply knowledge and analyze problems as a prelude to thinking about alternative solutions. However, only the most able students can synthesize contrasting points of view in order to develop new theoretical formulations, to propose and defend alternative solutions to problems, or to evaluate critically alternative courses of action.

Several explanations might be offered to account for the disparity between high- and low-achieving classes in the instructional goals teachers have been found to pursue. Metz (1978) has argued that teachers resort to basic skills instruction in classes serving low-achieving students as a classroom management strategy. In this view, the routine tasks and slow pace of work that accompany much basic skills instruction keep low-achieving students busy while accommodating their preferences for easy work.

Neo-Marxists and critical theorists, on the other hand, see the disparity in instruction across academic tracks as flowing from the educational system's role in reproducing social inequality. In this view, tracking and ability grouping sort students on the basis of social class and ethnic background and then provide them instruction consistent with their future occupational destinies. The notion that low-track students are incapable of critical reasoning prepares these students to accept subordinate social roles while the presumed superior cognitive powers of the high-track students justify the more prestigious and powerful social roles they will later occupy (Bowles and Gintis, 1976).

The pervasive influence of behaviorism in curriculum and instruction provides a final potential explanation for variation across US secondary classes in the pursuit of higher-order instructional goals. As Shepard (1991) demonstrates, behaviorist theories imply that students learn best when complex learning tasks are broken down into smaller parts that are learned sequentially. Only when the earlier, simple steps are mastered is the learner ready for more complex tasks requiring analysis, hypothesis testing, and critical evaluation. This view of learning may help account for the differentiation of teaching objectives by student ability. Low-achieving students may chronically experience lower-order instructional emphasis because educators view these students as "stuck" in the early phases of the learning process. In contrast, the higher-achieving students, having mastered the basic skills, may be viewed as prepared to handle more complex learning tasks.

Indeed, Peterson (1988:7) criticizes the traditional elementary school mathematics curriculum "as based on the assumption that computational skills must be learned before children are taught to solve even simple word problems." At the secondary level, the curriculum also reflects this sequential notion. For example, in mathematics, US students commonly take pre-algebra, then algebra, then geometry. It is common to view higher-order objectives as more appropriate later in the sequence: proofs are expected in geometry but not algebra; mathematical reasoning is more appropriate in algebra than pre-algebra. As students progress through the grades, only the more able students appear to keep up with the academic mathematics curriculum, so that most will fall away before they encounter truly higher-order instructional objectives.

The interpretations we have described are not mutually exclusive. Teachers might emphasize low-level objectives in low-track classes as a classroom management strategy even if the larger educational system functioned to reproduce inequality in adult status. And neither of these views precludes the possibility that the curriculum has been constructed to facilitate a behaviorist conception of learning in which mastering basic skills is a necessary prelude to higher-order reasoning. These perspectives, taken together, suggest two hypotheses:

> H1. The higher the academic track of a class, the more likely a teacher will be to report an emphasis on teaching for higher-order thinking in that class; and

> H2. The higher the grade level of a class, the more likely a teacher will be to report an emphasis on higher-order thinking in that class.

These effects of track and grade may not operate similarly in each of the major academic disciplines. As Stodolsky (1988) points out, disciplinary specializations in the US are founded on different epistemological bases. Therefore, there may be real differences in the degree of hierarchical curricular organization across the disciplines, with the result that the effects of track and grade will also vary. For example, the US mathematics curriculum tends to be sharply differentiated by track. College-bound students take a markedly different sequence of courses than do vocational or general track students. Honors students often take yet a different sequence. Moreover, the prescribed sequence of courses tends to be especially rigid in math (e.g., algebra before geometry before trigonometry). The science curriculum is similarly quite differentiated by academic track, and the prescribed sequence of courses fairly standardized (though not as standardized as in math) with biology preceding chemistry preceding physics. On the other hand, the social studies and English curricula appear less differentiated in both ways. Courses more often have the same labels for college-bound and non-college-bound students and the sequence of courses is more flexible.[1] These differences may be related to the different social functions served by the several disciplines. For example, it is common to view math as critical in gate-keeping decisions determining later occupational attainment, while social studies is essential in promoting political socialization. It would therefore be functional for the math curriculum to be highly differentiated by ability and age but dysfunctional for the social studies curriculum to be similarly differentiated.

The notion that the disciplines differ in their degree of hierarchical organization leads us to a third hypothesis:

> H3. The effects of track and grade level on teaching for higher-order thinking should be greater in math and science than in English and Social Studies.

In summary, our discussion to this point has focused on intra-teacher variation in teaching for higher-order thinking. We have conjectured that a given teacher intentionally differentiates instructional goals across the various classes he or she teaches and that when teaching younger students or students in lower academic tracks, a teacher tends to places less emphasis on higher-order instructional objectives. If such intra-teacher decision processes were the only source of variation in teaching for higher-order thinking, then variation among teachers would arise only to the extent that

---

[1]A preliminary analysis of the data used in this study confirms these ideas. We examined course titles by track designation, course titles by grade level, and the amount of ability grouping by subject area. These analyses lent support to the notion that math and science tend to be more rigidly sequenced and tracked than do social studies and English.

some teachers were assigned to teach more high-track or high-grade classes than were other teachers. We would then conclude that the effects of track and grade are deeply institutionalized in conceptions of teaching and learning that are essentially invariant across teachers and organizational environments.

Teacher Background and Training

An exclusive focus on intra-teacher processes neglects effects of relatively stable differences in the background and preparation of teachers on teaching for higher-order thinking. Yet these differences might be important in understanding the variation in instructional objectives that students encounter. There has been considerable commentary on the kinds of preparation and knowledge teachers need in order to teach effectively for higher-order thinking. Talbert and colleagues (1990) extensively reviewed this literature, focusing on two types of knowledge: in-depth knowledge of subject matter and "pedagogical content knowledge" (Shulman, 1987), that is, in-depth knowledge of how to teach the subject matter. Teachers who lack the required knowledge apparently resort to "teaching the facts." An emphasis on transmitting procedural and factual information allows teachers to control the flow of classroom interaction, thereby avoiding challenging discourse or responding to student challenges or questions that require a deep understanding of the subject matter being taught.

In this paper, we index teachers' broad subject-matter knowledge by reference to teachers' level of education, and we index broad pedagogical knowledge by reference to years of teaching experience, under the assumption that more highly educated and experienced teachers know their subject areas better and have acquired more pedagogical expertise. This reasc 3 generates two hypotheses:

> H4: The higher a teacher's level of education, the greater the likelihood that the teacher will emphasize teaching for higher-order thinking; and
>
> H5: The more years of teaching experience possessed by a teacher, the greater the likelihood that the teacher will emphasize teaching for higher-order thinking.

These hypotheses relate teacher background to a broad predisposition to teach for higher-order thinking. However, the knowledge required to teach for higher-order thinking is undoubtedly highly specific. For example Talbert et al. (1990) discuss situations in which teachers who are generally capable of teaching for higher-order thinking do not teach in this way because of a mismatch between their knowledge and the particular content of the course or lesson they are assigned to teach. In fact, Talbert et al. (1990) argue that "cross-over" teaching — teaching out of one's subject area -- strongly predicts "transmission" style instruction, which they view as the opposite of teaching for understanding. In this paper, we index the specific preparation of teachers by their reports of the degree to which they feel prepared to teach each of the courses to which they are assigned, suggesting the following hypothesis:

> H6: The more prepared a teacher feels to teach a particular course, the more likely that teacher will be to emphasize teaching for higher-order thinking in that course.

Our hypotheses regarding effects of teacher education and experience (H4 and H5) are hypotheses about predictors of <u>inter-teacher</u> variation in instructional objectives. In contrast, the hypothesized effects of the match between a teacher's preparation and the content to be taught in a particular class are, in part, <u>intra-teacher</u> effects. To the extent a teacher feels differentially well prepared to teach in the several classes assigned, that teacher's instructional goals are expected to vary across those classes. Of course, it might be that some teachers are, on average, better-prepared to teach their classes than are other teachers, contributing to inter-teacher variation in instructional goals.

Schools as Formal Organizations

Finally, recent research asserts a relationship between the formal organization of schools and the distribution of teaching for higher-order thinking. For example, Darling-Hammond and Wise (1985) argue that the work of teachers is becoming more routinized and regulated as state education agencies and local school systems increasingly implement standardized curricula and use standardized tests to assess the performance of students, teachers, and schools.

Available evidence suggests that teachers alter their instructional practices in response to these policies (Bullough, Gitlin, & Goldstein, 1984; Darling-Hammond and Wise, 1985; Rosenholz, 1987; for a review, see Rowan, 1990). Because available texts and tests tend to stress "low-level" cognitive skills, and because bureaucratic pressures for accountability force teachers to use "direct instruction" to teach to these tests, instruction is allegedly becoming more highly focused on basic skills. Moreover, teachers appear to alter the pace of instruction in response to bureaucratic policies, stressing content coverage over depth, a practice that Newmann (1990) has argued is incompatible with teaching for higher-order thinking.

These ideas suggest that two organizational features of schools will have direct effects on teachers' emphasis on higher-order thinking: the extent to which standardized tests of student achievement are used to judge the quality of teaching and the extent to which teachers feel obligated to "cover the curriculum." In our view, these organizational features will produce variation at two levels of our model. To some extent, accountability systems and pressures for curriculum coverage operate at the school level and should produce inter-school differences in teaching objectives. At the same time, however, previous research demonstrates that not all teachers will feel these pressures equally (Rowan, Raudenbush, & Kang, 1991). Teachers who inhabit different locations within the organization of the school may differentially perceive such pressures or may respond differently to them. We shall test the following hypotheses:

> H7: To the extent a teacher feels that the quality of his or her teaching is judged on the basis of standardized tests, that teacher will be less likely to teach for higher-order thinking; and
>
> H8: To the extent that a teacher feels pressure to "cover the curriculum," that teacher will be less likely to teach for higher-order thinking.

As we have seen, much of the debate about organizational influences on teaching for higher-order thinking has centered around the negative effects of school bureaucratization. However, an equally important line of work has emphasized the importance of school restructuring to promote teaching for higher-order thinking. In this argument, such teaching will require the development of a collegial "learning community" among teachers, supportive leadership from school administrators, and more control by teachers over the kinds of instructional policies and procedures that currently encourage a focus on low-level cognitive skills. As with bureaucratic controls, we expect that schools might systematically differ on these dimensions, but there is also strong evidence that teachers within the same school vary in the extent to which they receive administrative support, join with other teachers in collaborative relations, and obtain control over important instructional policies (Rowan, et al., 1991). Thus, as with hypotheses about bureaucratic controls, we add hypotheses about these organizational factors to predict inter-teacher variation:

> H9: The more a teacher receives support from school administrators, participates in collaborative relations with other teachers, and exercises control over key policies in the areas of curriculum and instruction, the more likely that teacher will be to emphasize teaching for higher-order thinking.

Data

Data for the study were collected in 16 high schools in California and Michigan. The 16 schools were chosen purposefully to guarantee diversity in secondary school teaching contexts in terms of state policies, district resources, school organization, and student composition. In the spring of 1990, questionnaires were administered to all teachers in these schools. Response rates varied from a low of 50% to a high of 100% with a median response rate of 75%. Mathematics, science, social studies, and English teachers were administered a specially tailored, subject-specific questionnaire that asked them to report information on each of the classes they were teaching. This procedure yielded complete data on 1205 classes taught by 303 academic teachers.

Characteristics of the Classes

Table 1a presents descriptive statistics on characteristics of these 1205 classes. The average class size was 25.69 students (sd = 6.47). Grade levels ranged from 1 = freshman to 4 = senior (m = 2.34; s.d. = 1.00). In the analyses below, we included class size as a control variable.

Insert Table 1 About Here

For each class, academic teachers were asked to classify the academic track of the class according to the following coding scheme: vocational, general, college-bound (non-honors), honors, and mixed. In this sample, only four percent of the classes were described as mixed, implying that teachers had little difficulty in classifying their classes according to track. Mixed classes were dropped from the analysis. For purposes of the analysis, two indicator variables were created: an "honors" indicator took on a value of one for honors classes, zero otherwise; and an "academic" indicator took on a value of one for college-bound (non-honors) classes, zero otherwise. The classes having values of zero on both indicators were the vocational and general track (non-college) classes. Table 1a shows that of the 1205 classes, 192 (16 %) were honors classes, 506 (42 %) were academic but not honors, and the remaining 507 classes (42 %) were either general or vocational track classes. Teachers were also asked to report their specific level of preparation to teach each class. This variable was recoded as an indicator variable that took on a value of one if a teacher felt "very well prepared" for that class, zero otherwise. In 61% of the classes, the teachers felt very well prepared.

Characteristics of Teachers and School Organization

Table 1b presents descriptive statistics for variables measured at the teacher level. The first set of variables describe the education and experience of teachers. Teachers in this sample were both highly experienced, with mean years of experience = 20.02 (s.d. = 9.15), and highly educated (70 % with masters degrees). We also collected data on sex and ethnic background: 91 % were white, and 61 % were male.

Three scales were constructed to measure the organizational environments in which teachers worked. As discussed above, we view these as teacher-level variables, in part because responses reflect teacher perceptions of characteristics of their schools; and in part because previous research has found that these measures vary within schools in response to teacher locations within the academic division of labor (Rowan et al., 1991). We note that global differences in school organization not measured here, as well as differences among schools in student composition and academic mission, are taken into account by including in the model a school-level variance component.

Bureaucratic pressures on teachers were measured by two Likert-type scales. Pressures arising from standardized testing were assessed by an item that asked teachers about the extent to which colleagues "judge the quality of my teaching on the basis of my students' achievement gains" (m = 3.97, sd = 1.27). Pressures for curriculum coverage were assessed by asking teachers the degree to which it was important to "cover the curriculum" (m = 4.51, sd = 1.15).

Variables measuring aspects of school organization hypothesized to promote teaching for higher-order thinking were measured by scales used in previous research (Pallas, 1988; Rowan, et al. 1991). Supportive principal leadership was measured by a 13-item scale with items indicating such diverse principal activities as effectively coping with outside pressures, setting priorities, recognizing, encouraging, and supporting staff, and involving staff in decision making. Internal consistency was .92. The extent of staff cooperation and collegiality was measured by a scale including six items indicating the extent to which staff members help each other in diverse duties, share beliefs and values about the central mission of the school, maintain uniformly high standards of performance for themselves, and seek new ideas. Internal consistency was .86. Teacher control is a nine-item scale indicating teacher control over student behavior codes, content of inservice programs, student grouping, school curriculum, text selection, teaching content and techniques, and amount of homework assigned. Internal consistency was .75. These measures show a close correspondence to ideas presented by Rosenholz (1985) and Little (1982) about the characteristics of effective school organization.

Measures of Teaching for Higher-Order Thinking

The outcome of interest in this paper is the amount of emphasis teachers place on higher-order thinking in each of the classes they teach. The term higher-order thinking has multiple intellectual roots, but in current usage it has come to denote a form of cognitive understanding constituted by knowing the general patterns and principles in a particular knowledge domain and by comprehending the relationships among these patterns and principles. Moreover, higher-order thinking has come to denote the application of such knowledge in problem-solving (Cole, 1991).

Although higher-order thinking undoubtedly shares common features across disciplines, we decided to treat the phenomenon as discipline-specific rather than generic. We therefore developed separate measures of teaching for higher-order thinking for mathematics, science, social studies, and English. In science and math, we used items developed as part of the NELS 90 survey, administered by the National Center for Education Statistics

By contrast, items for social studies and English are original in the current study.

Because the development of measures of teaching for higher-order thinking was, in itself, informative about instruction in high schools, we report on this measurement work in some detail. In each of the four questionnaires, we used a set of eight discipline-specific items that asked teachers to rate the degree of emphasis they placed on particular learning objectives for each of the classes they taught. Degree of emphasis on an objective was indicated by a four-point Likert scale including the choices "none", "a little", "moderate", and "heavy" emphasis.

The scale for each discipline included items intended to indicate both "higher-order" and "lower-order" objectives, where the latter involve tasks requiring memorization or the development of procedural skills. We expected that responses to items included would be sorted according to whether they indicated higher-order or lower-order objectives and that the two sets of responses would be quite highly negatively correlated.

As a next step in the analysis, we performed separate factor analyses on the discipline-specific item sets using classes as the unit of analysis. We assumed that the factor analysis would yield a single common factor for each discipline with higher- and lower-order items having opposite signs. However, in the math, science, and social studies areas we instead found that the factor analyses yielded two separate factors, one for higher-order objectives and one for lower-order objectives (the results for English are described below). The implication was that embracing higher-order objectives in math, science, and social studies classes did not preclude also emphasizing lower-order objectives.

After performing the factor analyses, we combined items from the identified factors into separate scales and assessed the reliabilities of these scales. This analysis indicated that the item sets tapping lower-order objectives in the three disciplines formed scales with weak reliability, with little variability in scale means across classes. In contrast, the item sets tapping higher-order objectives formed quite reliable scales with scale means varying quite substantially across classes. Based on these results we decided to utilize the higher-order scales in the analyses reported below but to drop the lower-order scales.

These measurement findings were consistent with previous reports about instructional emphasis in high school classes. We found a fairly pervasive emphasis on teaching lower-order cognitive skills in math, science, and social studies, a finding consistent with the observations of Goodlad (1984) and Powell, et al. (1985). On the other hand, we found variation across classes in the amount of emphasis placed on teaching for higher-order thinking. Thus, the measurement results indicate a uniform emphasis on basic skills and a varying emphasis across classes on teaching for higher-order thinking.

The measurement analysis yielded discipline-specific scales of the amount of emphasis teachers placed on higher-order thinking in each of their classes. In math and science, the items included in these scales corresponded closely to the descriptions of higher-order thinking advanced by Cole (1991) and Prawat (1989). The items indicate emphasis on key concepts, principles, and relationships in a discipline and an emphasis on the use of this knowledge in problem-solving situations.

Math. For math, the scale had four items. Items in the scale were standardized separately and then averaged to form a scale score. To ease interpretation, this initial score was then re-scaled to have a mean of 50 and a standard deviation of 10. The resulting scale had an internal consistency of .75 ($m = 49.68$, $sd = 9.72$). The items in the scale were: (1) understanding the logical structure of mathematics; (2) understanding the nature of proof; (3) knowing mathematical principles and algorithms; (4) thinking about what a problem means and ways it might be solved.

Science. The science scale consisted of five items which tended to place heavy emphasis on scientific problem-solving. The procedure for deriving the overall score for this scale was the same as for math. The science scale had an internal consistency of .79 ($m = 50.0$, $sd = 9.95$). The five items were: (1) prepare students to evaluate arguments based on scientific evidence; (2) teach scientific methods; (3) develop problem-solving/inquiry skills; (4) develop scientific writing skills; and (5) develop skills in lab techniques.

Social studies. The social studies scale developed here mirrors Newmann's (1990) discussion of higher-order thinking in high school social studies classes. In addition to understanding key concepts and principles in a discipline, Newmann placed emphasis on engaging students in challenging problems and having students manipulate information to solve these problems. Again, the overall scale score was derived in a manner similar to the derivation of the math and science scale scores, resulting in a four-item scale ($m = 49.53$, $sd = 9.90$) having internal consistency .76. The four items were: (1): formulating and presenting arguments to a group; (2) critically evaluating historical accounts or arguments; (3) analyzing historical and social science theories; and (4) using historical concepts to interpret current social issues.

English. The factor analysis of items for the English scale yielded results different from the results for the other disciplines. In part, the difference reflects the different approach to higher-order thinking developed by specialists in this area. Bereiter and Scardamalia (1987) present a discussion of higher-order thinking in the areas of reading and writing that builds on Resnick and Resnick's (1977) conception "high literacy." Bereiter and Scardamalia argue that higher-order thinking in reading and writing involves going beyond the extraction of meaning from simple texts (i.e., "lower" literacy). Instead, students are seen to be engaged in "high literacy" when they read literary classics, make inferences about these texts through writing and discussion, and when they attempt to gain personal meaning from the literature they are reading.

The factor analysis of items for the English scale yielded two separate factors, both of which are included in this study. An initial factor appears to denote a concept of "high literacy" as it applies to reading. This five-item scale ($m = 50.87$, $sd = 9.49$) had an internal consistency of .84. The items were: (1) writing about literature in a variety of forms (freewrite, formal essays, etc.); (2) comprehending plot and basic meaning of assigned literature; (3) defining literary terms; (4) engaging in literary analysis; and (5) exploring personal responses to a variety of literature. A second factor is specific to writing. This two-item scale ($m = 50.79$, $sd = 9.44$) had an internal consistency of .60. The two items in the scale were: (1) writing about literature in a variety of forms; and (2) learning to pre-write, draft, revise, and edit. Note that item (1) in the reading and writing scales are the same.

Partitioning Variation Among Classes, Among Teachers, and Among Schools

The measurement analysis indicated that, for each of the four disciplines studied, there is reliable class-to-class variation in the extent to which teachers emphasize higher-order learning objectives. We know that the 1205 classes under study vary in this way because only if such variation existed could the scales that measure such instructional objectives achieve reasonably high internal consistencies with modest numbers of items per scale. However, this reliable class-to-class variation could arise for three very different reasons, each having fundamentally different implications for understanding how teachers develop instructional goals.

First, classes could vary reliably because the same teacher differentiates her objectives across the different classes to which she is assigned. We call this reliable intra-teacher variation. Notice that it would be absurd to label teachers "higher-order teachers" or "lower-order teachers" if the main source of reliable variation among classes is within teachers. Rather, it would be crucial to understand the characteristics of classes that elicit higher-order objectives and to understand the basis of a teacher's differential judgments across classes regarding the appropriate objectives.

Second, there could be reliable variation across classes because there are important differences among teachers in emphasis on higher-order objectives. We call this reliable inter-teacher variation. In the extreme case, there would be no variation within teachers across classes, that is, a teacher would pursue exactly the same objectives in each of her classes. In this instance, we should waste no time studying how characteristics of the different classes to which a teacher is assigned influence that teacher's choice of objectives because the choice never varies. Rather, we should focus on differences among teachers that predict their pursuit of different objectives.

Third, it be might that reliable variation among classes arises because schools vary substantially in the objectives pursued by their teachers. We call this reliable <u>inter-school</u> variation. In the extreme case, every teacher in a school would pursue exactly the same objectives in every class so that any differences among objectives pursued for different classes are strictly determined by the school in which that class is located. In this instance it would be futile to study either class differences or teacher differences that predict variation within a school. Rather, one should concentrate on uncovering how different schools shape the objectives of each teacher within them.

Using methods described in detail in Raudenbush, Rowan, and Kang (1991), we decomposed the variance in the outcome measures into the three components described above. This analysis was based on a three-level hierarchical linear model using the computer program HLM3 (Raudenbush, Bryk, Seltzer, & Congdon, 1990). This program produces maximum likelihood estimates of the variance and covariance components based on unbalanced data. In the simplest three-level model, which we refer to as the "baseline" model, there are no predictor variables. Later, we shall utilize a three-level regression model in an attempt to predict the variation identified by the base model. Based on the variance decomposition from the baseline model, one can estimate the reliability of inter-teacher and inter-school differences in instructional objectives. We also can obtain a rough picture of the reliability of intra-teacher differences from this model. The results are illuminating, and vary substantially across the four disciplines.

Insert Table 2 About Here

<u>Decomposing the Variation Among Math Classes</u>

Table 2 shows that the intra-teacher variance in higher-order math objectives is estimated at 60.11, the inter-teacher variance at 33.60, and the inter-school variance at .09. Hence, 64.1 percent of the variance in higher-order objectives is within teachers, 35.8 percent is among teachers within schools, and a trivial proportion of the variance lies among schools. Two other statistics are helpful in interpreting these variance components. First, the Table presents chi-square tests of the null hypothesis that the true variance among teachers within schools is null, in which case the estimated variance of 33.60 is an artifact of chance. This hypothesis is easily rejected since the observed chi-square of 224.17 far surpasses the critical value at the .001 level based on 60 degrees of freedom. Similarly, we can test the null hypothesis that the true variance among schools is null. In this case we cannot reject this null hypothesis because values of an estimated variance near ours (.01) are highly plausible when the null hypothesis is true.

The table also provides information about the reliability of estimates at the teacher and school levels. The reliability for discriminating among teachers within a school is estimated to be .43, while the reliability for discriminating among schools is estimated to be a trivial .01. The three-level program does not produce estimates of the reliability of intra-teacher differences. However, we know that the internal consistency of the scale score for math classes is .75 (see Table 1), implying that about 24 % of the total variance in the scale scores is attributable to measurement error. This represents a variance of about 23.1; when compared to the intra-teacher variance estimate of 60.1, we see that about (60.1-23.1)/60.1 or 62% of the intra-teacher variance is reliable.

In sum, the majority of the variance on the math scale is intra-teacher variance with nearly all of the remaining variance being inter-teacher variance. There is little evidence of inter-school variance. Although the teacher-level variance is statistically significant and substantial in magnitude, differences among teachers are less reliable than are intra-teacher differences.[2]

These results encourage a search for differences among the classes to which a teacher is assigned that might predict the large component of intra-teacher variation. The data also show potential for using teacher differences to predict teacher variation within schools. The data provide little hope for a search for school effects on the math scale.

Decomposing the Variation Among Science Classes

The results for science are listed in the second column of Table 2. The estimated intra-teacher variance is 38.71; the estimated inter-teacher

---

[2]The reliability of the data for discriminating among teachers is limited by the number of classes per teacher, but this number results from the structure of secondary education and not from the design of the study. Including more items would improve the teacher-level reliability. The substantial variation within teachers across classes reduces the reliability of the data for understanding differences among teachers.

variance is 59.50; and the estimated inter-school variance is 1.37. Thus we estimate that 38.9 % of the total variance is intra-teacher variance, 59.8 % is inter-teacher variance, and only 1.4 % is inter-school variance. Thus, the results differ from the math results in that the majority of variation is at the teacher level, while a substantial fraction is within teachers. The results are like the math results in that little of the variation is among schools. Once again, the hypothesis that teachers within schools do not vary is easily rejected while the hypothesis that schools do not vary is readily retained.

Despite a large component of inter-teacher variance, the reliability for discriminating among teachers within schools is modest (see note 2) at .49. The inter-teacher variance clearly has a reliable component because the estimated intra-teacher variance of 38.71 is nearly double the estimated measurement error variance component of 20.6.

Decomposing the Variation Among Social Studies Classes

The results for social studies are listed in the third column of Table 2. The estimated intra-teacher variance is 19.24; the estimated inter-teacher variance is 63.71; and the estimated inter-school variance is 11.43. Thus, about 20.4 % of the variance is within teachers, 67.5 % is among teachers within schools, and 12.1 % is among schools. Thus, the majority of variation is at the teacher level, and, not surprisingly, the hypothesis that teachers within schools do not vary is easily rejected. In this case, the hypothesis that schools do not vary can be rejected at the marginal significance level of $p = .03$. Thus, there is some evidence of school-to-school variance for social studies.

Despite the large component of inter-teacher variance, the reliability for discriminating among teachers within schools remains a modest .53. The school-level reliability is estimated to be .29. The intra-teacher variance shows less evidence of having a reliable component than for math and science because the intra-teacher variance estimate is about the same as the estimated measurement error variance component.

Decomposing the Variation Among English Classes

Recall that, for English, two scales emerged, one for literary emphasis and one for emphasis on writing. We consider each below.

Literary emphasis. The fourth column of Table 2 lists the variance estimates, hypothesis tests, and reliabilities for literary emphasis. In sum, we found that 50.8 % of the variance was within teachers; 35.5 % among teachers within schools, and 13.7 % among schools. Both the inter-teacher and inter-school variances were significantly greater than zero. Strong evidence existed that a reliable component of variation existed within teachers.

Writing emphasis. We found that 35.4 % of the variation was within teachers; 46.7 % was among teachers within schools, and 17.9 % was among schools. Both the inter-teacher inter-school variances were significantly greater than zero. Evidence that of reliable intra-teacher variance was weak.

Summary

Decomposition of variance in these scales is important because it has strong implications for where one might look to explain why classes vary in terms of the instructional objectives teachers pursue. In this regard, we discovered important disciplinary differences. There was little evidence of inter-school variance in math or science objectives, while evidence of inter-school variance was manifest for both social studies and English objectives. Large and reliable components of intra-teacher variance were found for math and science, while inter-teacher variance was more important in social studies. In English the results were a bit more complicated with most variation in literary objectives found to lie within teachers and most of the variation in writing objectives found to lie among teachers. Having identified the components of variance in each discipline, the next task of the analysis was to employ a three-level regression model to predict this variation.

Correlates of Teaching for Higher-Order Thinking

Approach

For each discipline, multilevel regression models were fitted in three stages. First, predictors that vary within teachers were specified in an intra-teacher prediction model. These included the two indicators for student track (the academic and honors indicators), student grade, and teacher preparation.[3] The analysis also controlled for class size. Specifying the intra-teacher predictors enabled us to estimate the adjusted intra-teacher and inter-teacher variance, that is, the amount of variance within and among teachers after controlling for the effects of the classes to which they were assigned.

In the next stage, teacher background variables, including years of experience, sex, and race, were added and their effects estimated in an initial inter-teacher model. Control for such background characteristics is important to insure that estimates of hypothesized effects were not biased

[3]Recall that teacher preparation varied from class to class because we asked each teacher to report her level of preparation to teach the particular subject matter in each class.

by uncontrolled background characteristics. Predictors with coefficients less than 1.5 times their estimated standard errors were tentatively dropped from the model.

At the third stage, the inter-teacher predictors hypothesized to influence instructional objectives were added. These included teachers' level of education, emphasis on covering the curriculum, emphasis on achievement test gains, and perceptions of principal leadership, staff collaboration, and teacher control. Predictors with coefficients found to be less than 1.5 times their estimated standard errors were dropped and the model re-estimated. To guarantee that no predictor, including teacher background characteristics, could have been mistakenly dropped from the model, residuals were regressed on variables excluded from the model. If the approximate "t-to-enter" (see Raudenbush, Bryk, Seltzer, & Congdon, 1990) exceeded 1.5, that predictor was re-entered into the analysis. Hence, for each outcome we arrived at a final model including predictors related to the outcome and excluding predictors unrelated to the outcome. We used the relatively liberal criterion of $t = 1.5$ because we worried more about the bias that might arise from failing to specify a predictor than about the lack of efficiency that arises when the model is slightly over-fit.

After estimating the final model for each outcome, we assessed the amount of inter-teacher variance explained. This is the difference between the adjusted inter-teacher variance (based on the intra-teacher model) and the residual inter-teacher variance (based on the final model) divided by the adjusted teacher level variance.

With only 16 schools, the sample offered only limited opportunities to model the variation among schools in that identifying and controlling for relevant covariates and theoretically important predictors would be of questionable validity with only 15 degrees of freedom. This shortcoming is unimportant in the case of mathematics and science given lack of evidence that schools varied on the outcome.

The results for the final models in each discipline are presented in Table 3. Potential predictors are listed in the first column. Columns 2 - 6 provide results for math, science, social studies, literature, and writing. If a predictor was not related to an outcome, the table leaves that entry blank. The results for math and science were quite different from the results for social studies and English (literature and writing). We therefore describe these two sets of results separately.

Insert Table 3 About Here

Results for Mathematics and Science

Effects of track and grade. The crucial results for mathematics and science are the truly massive effects of track. In math, the degree of emphasis on higher-order objectives is, on average, 10.50 points higher for academic classes than for non-academic classes ($t = 11.34$). The gap between honors classes and non-academic classes is 17.44 points ($t = 14.52$). Recalling that the standard deviation of the math higher-order teaching scale is slightly less than 10 points, we see that the academic-nonacademic gap exceeds a standard deviation in magnitude and the gap between honors and non-academic classes exceeds 1.7 standard deviation units.

Large track effects appear in science as well. The academic-non-academic gap is 10.25 ($t = 11.30$), though in this case the honors-non-academic gap is similar at 12.53 points ($t = 10.10$).

There are some differences between the math and science results. In math, a significant effect of grade appears, $b = 1.38$, $t = 3.41$. This result implies, for example, that the degree of emphasis on higher-order objectives would be 3*1.38 = 4.14 points higher for seniors than for freshmen, an effect slightly in excess of .41 standard deviation units. The grade effect in science is in the same direction, but is smaller and of marginal statistical significance, $b = .78$, $t = 1.66$, implying an expected gap between seniors and freshmen just exceeding 3*.78 = .23 standard deviation units.

Teacher preparation and organizational effects. Results for the two outcomes are quite similar with regard to teacher preparation and organizational effects. No effect of level of preparation was manifest for either outcome, and in science, no teacher education effect appeared. In math, the teacher education effect was in a direction opposite to that predicted, $b = -4.06$, $t = -2.19$. No organizational effects appeared in science. In math, a small tendency for teachers emphasizing curriculum coverage to also emphasize higher-order objectives appeared, but failed to achieve conventional levels of significance, $b = 1.22$, $t = 1.62$ $p > .10$. In sum, the analysis revealed no support for the explanations based on teacher preparation and organizational design effects for mathematics or science.

Explanatory power of the model. The model accounted for an impressive 59.6 % of the within-teacher variance for math and 47.6 % for science (Table 4). Given the level of measurement error variance discussed earlier, these statistics quite substantially under-report the explanatory power of the model at this level.

Insert Table 4 About Here

Neither model explains substantial inter-teacher variance. For science, none of the variance is explained because no teacher-level predictors were significantly related to the outcome. For math, the model accounted for about seven percent of the inter-teacher variance.

Results for Social Studies and English

Effects of track and grade. Substantial track effects appeared in all three areas (social studies, literature, and writing), but the effects were not as massive as in math and science. For example, in social studies, the gap between academic and non-academic classes in higher-order emphasis was 3.56 points ($t = 3.10$) and the gap between honors and non-academic classes was 9.07 points ($t = 8.06$). Again the standard deviation is slightly below 10 points so these effects can be interpreted as just exceeding .36 and .91 standard deviation units, respectively. Patterns for writing and literature are quite similar. No positive effect of grade appeared in any area, though unexpectedly, negative effects of grade appeared in the two English subjects.

Effects of teacher preparation and education. Significantly positive effects of teacher level of preparation appeared for social studies ($b = 2.74$, $t = 3.01$) and for literature ($b = 3.57$, $t = 3.40$). These represent expected gaps in higher-order emphasis between classes for which teachers feel "very well prepared" and other classes. No such effect appeared for writing. No effects of the overall level of teacher education appeared in any of the areas, implying that no effect of having earned a masters degree was present.

Effects of organizational design. Favorable effects of supportive principal leadership appeared in the two English subjects ($b = 1.52$, $t = 1.71$ in literature; $b = 3.14$, $t = 2.75$ in writing;) but not in social studies. A favorable teacher control effect appeared in social studies ($b = 3.91$, $t = 2.29$), but not English. To gauge the magnitudes of these effects, standardized coefficients ("betas") may be computed as beta $= bS_x/S_y$, where S_x and S_y are the standard deviations of predictor and outcome given in Table 1. Applying these formulas leads, for example, to standardized estimates of .24 for the effect of principal leadership in writing and .23 for the effect of teacher control in social studies. No effect of staff collaboration was found in any subject area.

Recall that we had hypothesized that emphases on curriculum coverage and evaluating teaching in terms of achievement gains would be linked to a decreased emphasis on higher-order objectives. No evidence emerged to support these hypotheses. In fact, an emphasis on achievement gains was positively associated with teaching for higher-order thinking in social studies ($b = 2.46$, $t = 3.09$); and an emphasis on curriculum coverage was positively related to such teaching for higher-order thinking in literature ($b = 2.30$, $t = 4.39$).

Explanatory power of the models. As compared to math and science, the models for social studies, literature, and writing were less powerful in explaining intra-teacher variance and more powerful in explaining inter-teacher variance (Table 4). This is not surprising. Track effects were not as massive in the three humanities as they were in math and science, while teacher and organizational predictors were more helpful in the humanities than in math or science.

Percentages of intra-teacher variance explained were 26.5, 15.1, and 19.7 for social studies, literature, and writing, respectively. The corresponding percentages of inter-teacher variance explained were 25.3, 40.6, and 13.5.

Discussion

A fundamental contradiction between current educational theory and practice provides the motivation for the present inquiry. On the one hand, an emerging consensus among educational theorists is that fostering higher-order thinking ought to be a principal goal of instruction in all academic subjects for all students. In contrast, field workers observing teaching practice in the US, both at the primary and secondary levels, have described teacher-student interactions that focus primarily on empirical and procedural issues and only rarely on higher-order thinking including the creative application of fundamental principles to novel situations, the criticism and synthesis of contrasting viewpoints, or the construction of alternative explanations. Ethnographic research cited earlier suggests that teaching for higher-order thinking, when it does occur, is far more likely to occur in high-ability classes than in low-ability classes, though theorists claim that the pursuit of such objectives is particularly important for students with limited academic background.

Measuring Emphasis on Higher-Order Objectives

We sought to evaluate plausible alternative explanations for variation in the pursuit of higher-order instructional objectives. This evaluation required, first, accurate measurement of the degree of emphasis on higher-order objectives in each of four subject areas. What we learned about this measurement problem may be of value to future inquiry on this important topic. The items designed to tap emphasis on higher-order objectives, adapted from the National Educational Longitudinal Survey of 1990, proved highly useful. A typical result is that a scale consisting of just four items achieves an internal consistency of about .75 across all classes. On the other hand, items designed to tap lower-order emphases, including memorization of facts and practice on procedures, were not helpful. Rather than loading negatively on the higher-order scale, they constituted a second, weakly reliable factor. This pattern was remarkably convergent across the four disciplines. Apparently, the pursuit of lower-order objectives does not preclude the simultaneous pursuit of higher-order objectives, at least, not in the minds of the teachers in this study.

The implications for future measurement efforts seem clear. It would be wise to write more items tapping the higher-order dimension for each subject area. Using variance component estimates from internal consistency analysis, we can estimate the expected internal consistency of each scale as a function of the number of items in that scale. For example, if the math scale consisted of 10 items rather than four, we estimate that the internal consistency of the scale would rise from .75 to around .88; a 15 item scale ought to produce an internal consistency of about .92. Having more items might also enable the measurement of interesting sub-dimensions of teaching for higher-order thinking, including, for example, a sub-scale that taps emphasis on higher-order cognitive content and a sub-scale that emphasizes classroom interactions that facilitate higher-order discourse.

Evaluating Alternative Explanations

We also considered three explanations for the under-emphasis on teaching for higher-order thinking. First, we reasoned that conceptions of teaching and learning rooted in the secondary curriculum may encourage teaching for higher-order thinking only in high-track classes and more advanced courses in a hierarchically ordered sequence. We predicted that these effects would vary across disciplines because the curriculum in these disciplines vary in the degree to which they are hierarchically organized according to track and grade. Second, we argued that inadequate teacher preparation might prompt a "transmission style" of teaching that avoids challenging teacher-student interactions. Such inadequate preparation could arise because the overall level of teacher education is inadequate or because of a poor match between teacher background and particular class assignments. Third, we hypothesized that school organizational structures put in place to emphasize direct instruction for basic skills may suppress teachers' adoption of higher-order objectives. On the other hand, organizational environments characterized by supportive administrative leadership, staff collaboration, and teacher control over the conditions of instruction were hypothesized to facilitate the pursuit of higher-order instructional goals.

The most prominent result of our study was the powerful link between track and emphasis on higher-order objectives. This link was strong and highly significant statistically in every discipline, but was especially pronounced in math and science. The second central result was that the correlates of teaching for higher-order thinking did depend upon the discipline. Hypothesized effects of grade appeared only in math and science.

No hypothesized effects of teacher preparation appeared in math and science, but important positive effects of teacher preparation did appear in social studies and literature. No organizational effects appeared in math and science, but some evidence of such effects were manifest in social science and on both English scales (writing and literature).

The teacher preparation effects arising in social studies and literature were effects of the teacher's self-reported preparation to teach a particular class. They were not effects of the overall level of teacher education. Our results provided no evidence that simply having obtained a master's degree predisposes a teacher to pursue higher-order objectives. Rather, the match between the teacher's preparation and the subject matter of a particular class appeared to be linked to higher-order emphasis.

The organizational effects were more complicated than anticipated. Some evidence of positive effects of supportive leadership and teacher control appeared; supportive principal leadership predicted an emphasis on higher-order objectives in writing and literature, and teacher control over the conditions of instruction predicted higher-order objectives in social studies. However, emphasis on curriculum coverage and on accountability for achievement gains did not relate to instructional objectives as hypothesized. In fact, rather than predicting low emphasis on higher-order objectives, emphasis on curriculum coverage predicted a higher-order emphasis in literature, and concern with achievement gains predicted a higher-order emphasis in social studies.

<u>Implications for Policy and Future Research</u>

<u>Mathematics and science</u>. The findings suggest that differentiated instructional objectives are strongly institutionalized in the secondary

mathematics and science curricula. In those disciplines, the typical teacher sets very different objectives for non-college and college classes, and there is evidence, especially in mathematics but suggestive also in science, that such a teacher sets higher-order objectives for courses taken later in a prescribed sequence. The grade effect in mathematics is not surprising: The secondary math curriculum is different from the other curricula in that the sequence of courses is to a larger extent hierarchical; certain courses (e.g., calculus) cannot be taken before certain other courses (e.g., algebra).

Commentary cited at the beginning of this paper rejects the contention that higher-order objectives are less appropriate for low-track classes or more elementary classes than for high track classes or more advanced classes. However, our study provides quite strong evidence that a diverse set of secondary math and science teachers disagree. If promoting higher-order thinking in secondary math and science is an important goal for educational policy, the advocates of such a policy must recognize that formidable institutional obstacles confront them. Formulas like providing more preservice teacher education, improving the match between teacher knowledge and teaching assignment, de-emphasizing standardized tests, or implementing school organizational reforms are by themselves or together unlikely to raise the conceptual level of discourse in the secondary math and science classes we studied. This does not mean that reforms of preservice education or school organization cannot be helpful. But such reforms must apparently challenge widely held conceptions of teaching and learning rooted in a curriculum that seems to have invariant characteristics across diverse secondary schools.

One might challenge these findings by criticizing the measures of teacher preparation and organizational environment employed in this study. One might especially challenge the absence of a large sample of schools that would have facilitated an exploration of school to school differences. While we certainly encourage researchers to work on improved teacher and school measures, our findings for math and science are strengthened by results of our variance decomposition. In both math and science, large components of variation are <u>within-teachers</u>. Such variation cannot be accounted for by the main effects of teacher-level characteristics, no matter how well measured. Moreover, substantial proportions of this large intra-teacher variance are accounted for by track. So our finding regarding track cannot be challenged by studies that improve measurement of teacher characteristics.

Moreover, no evidence emerged of important school-to-school differences in higher-order emphasis in math and science. So careful measurement of school-level characteristics, though certainly a laudable goal, would not likely be of much assistance in understanding the differentiation of objectives in these data. A critic might argue that idiosyncracies in our sample may have worked against finding school differences. However, we were able to find evidence of significant school differences in the other disciplines. The school-to-school invariance we report is specific to math and science. This discipline-specificity of findings is in itself an important result requiring further exploration.

<u>Social studies and English</u>. Important track effects appear in these disciplines as well, though they are not as powerful as in math and science. As in the case of math and science, this finding should encourage reformers to consider the institutional bases of the differentiation of instructional

objectives across classes. Widespread conceptions about the relevance of higher-order objectives for different kinds of students, arguably rooted in the secondary curriculum, should not be ignored.

However, the appearance of teacher preparation effects and school organization effects provides some hope to policy-makers seeking to increase the prevalence of higher-order instruction through reforms of preservice education and school organization alone. The capacity of a teacher to strive for higher-order objectives in each class appears to be undermined by mismatches between prior training or interest and specific class assignments. And these findings encourage further exploration of how supportive organizational environments might raise the level of discourse.

Critics might again argue that better measurement of teacher characteristics would shed more light on how improved teacher preparation, selection, and assignment might increase teaching for higher-order thinking; and that improved measurement of organizational characteristics might uncover important insights about the likely effects of organizational reform. In the case of social studies and English, we would agree. Our results show large components of inter-teacher and inter-school variation in the humanities. Though these findings do not undermine the importance of track, they do encourage the specification of better theories and measures to understand how teacher and school differences influence the pursuit of teaching for higher-order thinking in social studies and English.

Table 1 Descriptive Statistics for a) Class-Level and b) Teacher-Level Variables

a) Class-Level Variables

Variables	Coding and Range	Mean	St. Dev.
1. Class size[a]	(2, 38)	25.69	6.47
2. Student Grade Level[b]	(1, 4)	2.34	1.00
3. Honors Track vs Non-honor track	0 = non-honors track 1 = advanced or honors	.16	.37
4. Academic Track vs Non-academic Track	0 = non-academic track 1 = academic track	.42	.49
5. Level of Preparation	0 = not very well-prepared 1 = very well prepared	.61	.49
6. Emphasis on Literary Appreciation in English Classes 5 items[c] Reliability = .84	(16.43, 60.97)	50.87	9.49
7. Emphasis on Writing in English Classes 2 items[d] Reliability = .60	(15.71, 59.46)	50.79	9.44
8. Emphasis on Higher-Order Thinking in Social Studies 4 items[e] Reliability = .76	(29.22, 66.60)	49.53	9.90
9. Emphasis on Higher-order Thinking in Mathematics 4 items[f] Reliability = .75	(26.93, 63.89)	49.68	9.72
10. Emphasis on Higher order thinking in Science 5 items[g] Reliability = .79	(22.91, 65.60)	50.05	9.95

Notes to Table 1a

[a]Class size was centered around the grand mean in the HLM analysis.

[b]Student grade level was centered around the grand mean in the HLM analysis.

[c]There are 89 English teachers and 323 English Classes.

[d]There are 65 social studies teachers and 272 social studies classes.

[f]There are 74 Mathematics teachers and 304 Mathematics classes.

[g]There are 75 science teachers and 306 classes.

b) Teacher-Level Variables

Variables	Coding and Range	Mean	St. Dev.
1. Years of Experience[a]	(1, 40)	20.02	9.15
2. Level of Education	0 = no master's degree 1 = has master's degree	.70	.46
3. Race	0 = others 1 = whites	.91	.29
4. Sex	0 = female 1 = male	.61	.49
5. Staff Cooperation[b] 6 items reliability = .86	(-2.51, 1.49)	-.04	.74
6. Teacher control[c] 9 items reliability = .75	(-2.02, 1.22)	-.02	.58
7. Principal Leadership[d] 13 items reliability = .92	(-2.33, 1.34)	-.02	.71
8. Teacher Learning[e] 3 items reliability = .71	(-2.35, 1.53)	-.06	.76
9. Importance of Covering the Curriculum	(1, 6)	4.51	1.15
10. Achievement Gains Basis for Judging Teaching	(1, 6)	3.97	1.27

Notes to Table 1b

[a]Years of teaching was centered around the grand mean in the HLM analysis.
[b]Staff cooperation was centered around the grand mean in the HLM analysis.
[c]Teacher control was centered around the grand mean in the HLM analysis.
[d]Principal leadership was centered around the grand mean in the HLM analysis.
[e]Teaching learning was centered around the grand mean in the HLM analysis.

Table 2: Decomposition of Variance (no predictors specified)

	Math	Science	Social Studies	Literary	Writing
Intra-Teacher Variance	60.11	38.71	19.24	49.24	33.84
Inter-Teacher Variance	33.60	59.50	63.71	34.46	44.62
Inter-School Variance	0.09	1.37	11.43	13.29	17.16
<u>Reliability Estimates</u>					
Inter-teacher	0.43	0.49	0.53	0.41	0.45
Inter-school	0.01	0.06	0.29	0.36	0.37
<u>Components of Variance Tests</u>					
Inter-teacher					
Chi-square Statistic	224.17	409.28	576.04	267.02	373.09
df	60	60	50	73	73
p-value	0.01	0.00	0.00	0.00	0.00
Inter-School					
Chi-square Statistic	10.02	16.45	25.04	28.42	34.71
df	13	14	14	15	15
p-value	<.500	0.29	0.03	0.02	0.00

Table 3: Predictors of Emphasis on HOT Objectives

(a) Regression coefficient estimates
(with standard error estimates in parentheses)

	Math	Science	Social Studies	Literary	Writing
<u>Teacher-level predictors</u>					
Intercept	43.09	44.42	47.67	44.71	46.15
	(1.91)	(1.32)	(1.33)	(1.57)	(1.76)
Experience			-0.22		
			(0.11)		
Education	-4.06				
	(1.85)				
Stan. Test			2.46		
			(0.79)		
Coverage	1.22			2.30	
	(0.75)			(0.52)	
T. Control			3.91		
			(1.71)		
Staff Coop.					
Prin. Lead				1.52	3.14
				(0.89)	(1.11)
<u>Class-level predictors</u>					
Class size	0.14	0.02	0.04	0.23	0.11
	(0.07)	(0.06)	(0.06)	(0.08)	(0.07)
Grade	1.38	0.78	0.46	-0.80	-1.10
	(0.41)	(0.47)	(0.44)	(0.44)	(0.40)
Academic	10.50	10.25	3.56	5.27	4.26
	(0.93)	(0.91)	(1.15)	(1.20)	(1.23)
Honors	17.[illegible]4	12.53	9.07	8.95	8.69
	(1.20)	(1.24)	(1.13)	(1.23)	(1.16)
Level of Preparation	1.78	-0.87	2.74	3.57	-0.15
	(1.17)	(1.05)	(0.91)	(1.05)	(1.06)

Table 3 (Continued)

b) Variance components estimates

	Math	Science	Social Studies	Literary	Writing
Intra-Teacher Variance	24.29	20.27	14.14	41.81	27.19
Inter-Teacher Variance	35.35	64.65	43.88	13.75	38.54
Inter School Variance	0.13	2.56	7.31	17.30	23.40

Table 4 Variances and Variance Accounted For at the Class and Teacher Levels

	No predictors (from Table 2)	Class-level predictors only	Class-level & teacher-level predictors (from Table 3)	Variance Accounted for[a]
Math				
Intra-teacher	60.11	24.30	24.29	59.7%
Inter-teacher	33.60	39.47	35.35	10.4%
Inter-school	0.09	0.14	0.13	
Science				
Intra-teacher	38.71	20.27	20.27	47.6%
Inter-teacher	59.50	64.65	64.65	0.0%
Inter-school	1.37	2.56	2.56	
Social Studies				
Intra-teacher	19.24	14.16	14.14	26.5%
Inter-teacher	63.71	58.76	43.88	25.3%
Inter-school	11.43	10.36	7.31	
Literature				
Intra-teacher	49.24	41.61	41.81	15.1%
Inter-teacher	34.46	23.15	13.75	40.6%
Inter-school	13.29	14.53	17.30	
Writing				
Intra-teacher	33.84	27.12	27.19	19.7%
Inter-teacher	44.62	44.56	38.54	13.5%
Inter-school	17.16	19.97	23.40	

Note

[a] Percentage of intra-teacher variance accounted for = (Variance with no predictors - variance with all predictors) divided by the variance with no preditors.

Percentage of inter-teacher variance accounted for = (Variance with class-level predictors - variance with all predictors) divided by the variance with class-level predictors.

References

Bereiter, C. & Scardamalia, M. (1987). An attainable version of high literacy: Approaches to teaching higher order thinking skills in reading and wiring. Curriculum Inquiry, 17, 9-30.

Berliner, D. (1976). Impediments to the study of teacher effectiveness. Journal of Teacher Education, 27, 5-13.

Bowles, S. & Gintis, H. (1976). Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life. New York: Basic Books.

Bullough, R.V., Gitlin, A.D., & Goldstein, S.L. (1984). Ideology, teacher role, and resistance. Teachers College Record, 86, 339-358.

Cole, N.S. (1991). Conceptions of educational achievement. Educational Researcher, 19, 2-7.

Darling-Hammond, L. & Wise, A.E. (1985). Beyond standardization: State standards and school improvement. Elementary School Journal, 85, 315-335.

Doyle, K. (1983). Evaluating teaching. Lexington, Mass: Lexington Books.

Gage, N.L. (1979). The generality of dimensions of teaching. In P.L. Peterson and H.L. Walberg (Eds.), Research on Teaching: Concepts, Findings, and Implications. Berkeley, CA: McCutchan, pp 264-288.

Goodlad, J.L. (1984). A Place Called School. New York: McGraw-Hill.

Grossman, G. P. (1987). Teachers' political knowledge and attitudes. Social Studies Teacher, 8,4,7.

Hargreaves, D. H. (1967). Social Relations in a Secondary School. New York: Humanities Press.

Little, J.W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. American Educational Research Journal, 19, 325-340.

Metz, M. H. (1978). Classrooms and Corridors: The Crisis of Authority in Desegregated Secondary Schools. Berkeley: University of California Press.

Newmann, F. M. (1990). Higher-order thinking in teaching social studies: A rationale for the assessment of classroom thoughtfulness. Journal of Curriculum Studies, 22,1,41-56.

Oakes, J. (1985). Keeping Track: How Schools Structure Inequality. New Haven: Yale University Press.

Page, R.N. (1990). The lower track curriculum in a college-preparatory high school. Curriculum Inquiry, 20, 249-282.

Pallas, A. M. (1988). School climate in American high schools. Teachers College Record, 89,4,541-554.

Peterson, P. L. (1988). Teaching for higher-order thinking in mathematics: The challenge for the next decade. Perspectives on Research on Effective Mathematics Teaching (Volume 1, 2-26. Reston, VA: Lawrence Erlbaum Associates.

Powell, A. G., Farrar, A. E. & Cohen, D. K. (1985). The Shopping Mall High School: Winners and Losers in the Educational Marketplace. Boston: Houghton Mifflin.

Prawat, R.S. (1989). Teaching for understanding: Three key attributes. Teaching and Teacher Education, 5,315-328.

Raudenbush, S.W., Bryk, A.S., Seltzer, M. & Congdon, R. (1990). HLM 3.0: Computer Program and Users Manual. Unpublished manual. East Lansing: College of Education, Michigan State University.

Raudenbush, S.W., Rowan, B., & Kang, S.J. (1991). A multilevel, multivariate model for studying school climate in secondary schools with estimation via the EM algorithm. Journal of Educational Statistics, 16,4,295-330.

Resnick, D.P. & Resnick, L.B. (1977). The nature of literacy: An historical explanation. Harvard Educational Review, 47, 370-385.

Rosenbaum, J. E. (1976). Making Inequality: The Hidden Curriculum of High School Tracking. New York: Wiley.

Rosenholz, S.J. (1985). Effective schools: Interpreting the evidence. American Journal of Education, 93, 352-388.

Rowan, B. (1990). Applying conceptions of teaching to organizational reform. In R.F. Elmore and Associates (Eds.) Restructuring Schools: The Next Generation of Educational Reform. San Francisco: Jossey Bass.

Rowan, B., Raudenbush, S.W., & Kang, S.J.K. (1991). Organization design in high schools: A multilevel analysis. American Journal of Education. 99, 2, 238-266.

Shepard, L. (1991). Psychometricians' beliefs about learning. Educational Researcher, 20, 2-9.

Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1). 1-21.

Stodolsky, s. (1988). The Subject Matters: Classroom Activity in Math and Social Studies. Chicago: University of Chicago Press.

Talbert, J. E., Rowan, B., & McLaughlin, M. W. (1991). Understanding Context Effects on Secondary School Teaching. Unpublished manuscript, Palo Alto, CA: Stanford Center for Research on the Context of Secondary School Teaching.

US Department of Education, National Center for Education Statistics (1991). The Condition of Education, 1991, Volume I: Elementary and Secondary Education. Washington.